

LÉVY AREA WITH A DRIFT AS A RENORMALIZATION LIMIT OF MARKOV CHAINS ON PERIODIC GRAPHS

OLGA LOPUSANSCHI AND DAMIEN SIMON

ABSTRACT. A careful look at rough path topology applied to Brownian motion reveals new possible properties of the well-known Lévy area, in particular the presence of an intrinsic drift of this area. Using renormalization limit of Markov chains on periodic graphs, we present a construction of such a non-trivial drift and give an explicit formula for it. Several examples for which explicit computations are made are included.

CONTENTS

1. Introduction	2
1.1. General motivations	2
1.2. Structure of the present article	3
1.3. An easy discrete example made of rotating sums of Bernoulli r.v.	4
1.4. An example for comparing rough path and uniform topologies	5
2. The general framework	5
2.1. Periodic lattice and invariant processes	5
2.2. Decomposition into pseudo-excursions	7
2.3. Area process and rough paths	9
2.4. Embeddings of the discrete processes in continuous time	10
3. Convergence of the discrete process to the anomalous Brownian motion	10
3.1. Notations	10
3.2. Statement of the main theorem	10
3.3. Consequences of theorem 3.1.	11
3.4. Proof of the main theorem	11
4. Examples	17
4.1. The diamond model in dimension two	17
4.2. A three dimensional model with a non-trivial area anomaly.	19
5. Appendix	20
5.1. Rough paths	20
5.2. The group $G_2(E)$	21
5.3. The signed area	22
6. Open questions	23
References	24

1. INTRODUCTION

1.1. General motivations. Uniform convergence of paths is not the most useful topology for the study of the stochastic area of a process, as it can be seen in the following example ([8]). Consider the sequence of paths $\phi_n(t) = \left(\frac{\cos n^2 t}{n}, \frac{\sin n^2 t}{n} \right)$ in \mathbb{R}^2 for $t \in [0, 1]$. We clearly have $\|\phi_n\|_\infty \rightarrow 0$ when $n \rightarrow \infty$; however, at the same time, a simple computation of $A_n(t) = (\int_0^t \phi_n^1(s) d\phi_n^2(s) - \int_0^t \phi_n^2(s) d\phi_n^1(s))/2$ gives $\lim_{n \rightarrow \infty} A_n(t) = t/2$, so that clearly $\|A_n\|_\infty \not\rightarrow 0$. Actually, A_n represents the signed area between the graph of ϕ_n and $[\phi_n(0), \phi_n(t)]$. The uniform topology is too weak to encompass a complete description of the stochastic area. Similar problems arise in the study of sequences of differential equations driven by the ϕ_n s and lead to surprises regarding the convergence of the solutions.

We can palliate this problem by building out of the initial processes involved more complex structures, that allow to register all the relevant information. This is what does the theory of rough paths introduced by Terry Lyons. Loosely speaking, a (continuous) process in \mathbb{R}^d is considered as a first-level information, and we build the corresponding rough path by adding a few more levels. In particular, the stochastic area corresponds to the second level. The number of necessary levels is determined by the regularity of the process and each level is given by an iterated integral of a tensorial product (a double integral for level two, a triple one for level three and so on). A quick presentation of rough paths can be found in section 5.1. A more detailed introduction to this topics can be found in [12] and [15], and more exhaustive treatments in [8] or [7]. The rough path topology is then precisely a topology that makes solution maps of some differential equations continuous.

A very important example in stochastic analysis is the case of the *d-dimensional Brownian motion*, whose Hölder regularity is $\frac{1}{2} - \epsilon$ for any $\epsilon > 0$. Hence, we only need $\lfloor 2 + \epsilon \rfloor = 2$ (for small ϵ) components to construct the corresponding rough path. The corresponding rough path is thereby given by the process itself and its stochastic area, called the *Lévy area* (see section 5.3.2 for details). The second level of rough paths is made of a symmetric part and an antisymmetric part. The question of the definition of the symmetric part is related to the choice of stochastic integration (Itô or Stratonovitch for example). The antisymmetric part, the Lévy area, is not affected by this choice of the stochastic integral; the present paper focuses on the study of the choices that can affect this integral part. Thus, this is *not* a question of classical stochastic integration but this is a kind of *completion* of classical integration.

Since the Lévy area is an essential part of a Brownian motion rough path and it can influence the solution of an (S)DE, great interest is attached to studying the relation between it and the behaviour of the process it corresponds to. For instance, in [13], Antoine Lejay and Terry Lyons have studied how changing the area of a sequence of driving signals can influence its convergence and the solution of the limit (S)DE. In [11], Wilfrid Kendall introduces a method of coupling simultaneously two d -dimensional Brownian motions and the corresponding Lévy areas.

A special attention is given to Donsker-type theorems for rough paths. A strong theorem of Strook-Varadhan from [17] implies the convergence of the concatenation (product or sum depending on the group) of independent identically distributed variables to the Heisenberg Brownian motion in uniform topology. Emmanuel

Breuillard, Peter Friz and Martin Huesmann in [4] have strengthened this result, proving it still holds in rough path topology under additional moment conditions.

However, the area component itself is not given special attention in these examples, as it is regarded merely as an instrument for rough path construction. Besides, since the above examples, as many others of the same type, deal with sums of i.i.d. centred variables, the limit area (Lévy area) coincides with the area of the limit process (Brownian motion), which is a very particular case as we will see. This gives room to the following question: given a sequence of Markov processes that converges to the Brownian motion in uniform topology, what can the sequence of their areas tell us about the limit area? It is easy to see that it is not always exactly the Lévy area (see the example from section 1.3). A careful look at the rough path topology shows indeed that, *besides the choice of Itô or Stratonovitch stochastic integral*, there is some place left for additional terms in the limit Lévy area, such as a drift. We achieve the goal of building such anomalous drift in the area by replacing the random walks in [4] by suitable Markov chains.

Acknowledgements. D. S. thanks Q. Berger for valuable discussions on random walks on random environments at the origin of some constructions of the present papers.

1.2. Structure of the present article. The present article is organised into five sections, the introduction being the first one. The end of the introduction is devoted to the presentation of a very simple example which exhibits a non-zero area anomaly.

In section 2, we describe the general settings for our results: we define in particular Markov chains $(X_n)_{n \in \mathbb{N}}$ on periodic graphs and the stochastic signed area $(A_n(X))_{n \in \mathbb{N}}$ associated to these Markov chains and we present a decomposition of these processes which is based on excursion theory and inspired by renormalization ideas.

The first part of section 3 is dedicated to the proof of our main result (theorem 3.1) which is a generalisation to our class of Markov chains of the Donsker-type theorem for rough paths from [4]. We also discuss the consequences of this theorem on the universality class of the multidimensional Brownian motion and we express some caveats about the continuous description of the large size limit of discrete models.

The end of section 3 is devoted to the proof of the main theorem. In [5] and [10], authors have already considered the convergence of random walks on periodic graphs and Markov chains on graphs respectively, and in [3] an invariance principle has been proved for a certain class of random walks in random environment. In our case, we have a Markov chain on a periodic graph and we add the area component. We start by building out of the couple $(X_n, A_n(X))_{n \in \mathbb{N}}$ a continuous process by analogy with the classical Donsker theorem. Since the Markov chain $(X_n)_{n \in \mathbb{N}}$ takes values in a vector space E , the extended process $(X_n, A_n(X))_{n \in \mathbb{N}}$ and its continuous time version take values in the graded vector space $G_2(E) \subset E \oplus (E \otimes E)$ (see section 5 for definition and characterization of this group). We finally state and prove theorem 3.1, which establishes the convergence (up to an explicit multiplicative constant) of the continuous embedding defined previously to the anomalous Brownian motion $(B_t, \mathcal{A}_t + t\Gamma)_{t \in [0, \tau]}$ in the topology of $C^{0, \alpha-\text{Holder}}([0, \tau], G_2(E))$ for $\alpha < 1/2$, where Γ is a $d \times d$ antisymmetric matrix which represents the *area anomaly*.

In section 4, we present several examples for which we compute Γ by numerical simulations. Two of them (on a rectangular and on a hexagonal lattice respectively)

may be particularly interesting for physicists, as similar models arise in statistical mechanics.

Section 5 is dedicated to the definitions and generalities concerning the main objects we operate with in this article: rough paths, group $G_2(E)$, signed area etc. Since we present only aspects that are useful to our study, only a minimal overview of these subjects is given, and some references for further reading are proposed.

1.3. An easy discrete example made of rotating sums of Bernoulli r.v.

Let $(U_n)_{n \in \mathbb{N}^*}$ be a sequence of independent Bernoulli random variables such that $\mathbb{P}(U_1 = 1) = 1 - \mathbb{P}(U_1 = -1) = p$. We define two complex-valued processes $(Z_n)_{n \in \mathbb{N}}$ and $(Z'_n)_{n \in \mathbb{N}}$ in the following way: $(Z_n)_n$ is the random walk with increments chosen uniformly in $\{1, i, -1, -i\}$ and $(Z'_n)_n$ satisfies $Z'_0 = 0$ a.s. and, for $n \geq 1$, $Z'_n = \sum_{k=1}^n i^{k-1} U_k$. We set $X_n = \mathcal{R}(Z_n)$, respectively $X'_n = \mathcal{R}(Z'_n)$, and $Y_n = \mathcal{I}(Z_n)$, respectively $Y'_n = \mathcal{I}(Z'_n)$. A classical exercise in probability consists in checking that the laws of Z_n/\sqrt{n} and of $Z'_n/\sqrt{2np(1-p)}$ both converge to a normal law $\mathcal{N}(0, 1)$. Moreover, the whole processes $(X_n, Y_n)_n$ and $(X'_n, Y'_n)_n$ embedded in continuous time by linear interpolation converge in law to a standard Brownian motion in the uniform topology. The discrete Lévy area of the process $(X_n, Y_n)_n$ is defined as

$$A_n = \frac{1}{2} \sum_{1 \leq k < l \leq n} (X_k - X_{k-1})(Y_l - Y_{l-1}) - (Y_k - Y_{k-1})(X_l - X_{l-1})$$

and A'_n is defined in the same way for the second process. The process $(A_n/n)_n$ embedded in continuous time is known to converge to the Lévy area of the Brownian motion; the present paper deals with the rescaled Lévy area $(A'_n/(2np(1-p)))_n$ of the second process $(X'_n, Y'_n)_n$ and shows that, in the correct topology, it converges to the Lévy area of the Brownian motion with an *additional* drift γ . This drift is easily evaluated as:

$$(1) \quad \gamma = \lim_{n \rightarrow \infty} \frac{\mathbb{E}[A'_n]}{2np(1-p)} = \frac{(2p-1)^2}{8p(1-p)}$$

and some additional computations show that the limit of the first higher cumulants of $(A'_n/(2np(1-p)))_n$ coincide with the ones of the classical Lévy area.

Figure 1 describes the process (X'_n, Y'_n) as a Markov process in \mathbb{Z}^2 . Figure 2 presents histograms of the marginal laws of (X_n, Y_n, A_n) and (X'_n, Y'_n, A'_n) obtained by numerical simulations. This figure shows that, in the large n limit and with the classical rescalings, the two processes are very similar, except for the additional drift γ in the Lévy area, which we choose to call the *area anomaly*. Up to our knowledge, such a limit process in continuous time has never been described.

Intuition about the similarities and differences between the two processes can be quickly explained by the following renormalization argument. The increments of $(Z_n - Z_{n-1})_n$ are independent, whereas only the increments $(Z'_{4n+4} - Z'_{4n})_n$ are independent. In a time interval $\{4n, 4n+1, 4n+2, 4n+3, 4n+4\}$, the increments of (Z'_n) are bounded and thus do not contribute to the Brownian limit in the uniform topology; however, during the same time interval, correlations among these increments produce non-centered random areas. From a renormalization point of view, the local time correlations are irrelevant for the uniform topology but relevant for the Lévy area.

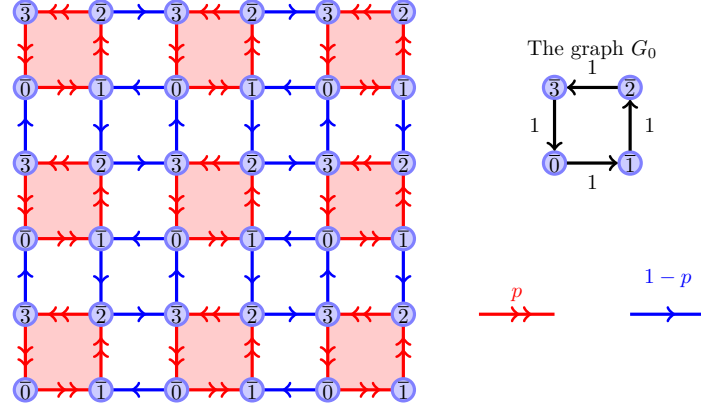


FIGURE 1. Markov chain $(Z'_n)_n$ of section 1.3. On the left hand side, its authorized transitions on the periodic graph $G = \mathbb{Z}^2$; on the top right hand side, its deterministic projection on the graph G_0 identified to $\mathbb{Z}/4\mathbb{Z}$. If $p \uparrow 1$, the chain tends to make more and more loops around the red shaded orbits.

1.4. **An example for comparing rough path and uniform topologies.** [from [6]] Let M be a matrix $n \times n$ such that the real component of its eigenvalues is strictly positive. Let B be an n -dimensional Brownian motion and $m > 0$ and consider the system

$$\begin{cases} dX = \frac{1}{m} P dt \\ dP = -\frac{1}{m} M P dt + dB \end{cases}$$

Then MX converges in the rough path topology to (W, \hat{W}) , where W is the Wiener process and \hat{W} is the Lévy area with a drift, which depends, in particular, on M . We would have gotten a Brownian motion without any dependence in M had we confined ourselves to the classical frame for a solution.

2. THE GENERAL FRAMEWORK

2.1. Periodic lattice and invariant processes.

Definition 2.1 (periodic subgraph). *Let E be a finite-dimensional vector space. A periodic subgraph of E is a infinite subset G of E such that:*

- (i) *all the points are separated,*
- (ii) *G is invariant under the translation action of a lattice $\Lambda \subset E$ on G .*

Property 2.1. *The graph can be decomposed as $G = \bigsqcup_{\lambda \in \Lambda} \lambda.G_0$ where G_0 is a finite subset of G and $\lambda.G_0$ is the translation of G_0 by λ .*

This property means that any point x of G can be parametrized in a unique way as (λ, x_0) where $\lambda \in \Lambda$ and $x_0 \in G_0$. We write $\lambda = \pi_\Lambda(x)$ and $x_0 = \pi_0(x)$ the two projections. We use alternatively the notation x or (λ, x_0) for a point in G .

Definition 2.2 (invariant Markov chain on G). *Let G be a periodic subgraph of E . A G -valued Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition law Q on a probability*

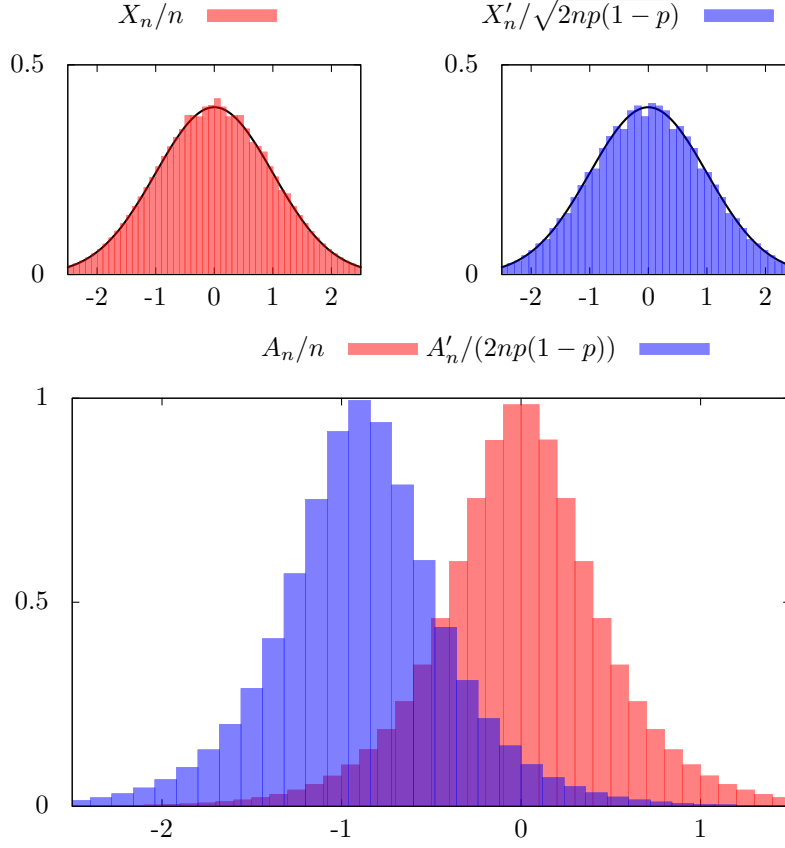


FIGURE 2. Empirical distributions of X_n/\sqrt{n} (top left), $X'_n/\sqrt{2np(1-p)}$ (top right), with the expected normal laws, and empirical distributions of the Lévy areas (below) A_n/n (red, right) and $A'_n/(2np(1-p))$ (blue, left). The parameters are chosen as $n = 250000$ and $p = 0.9$. Data are accumulated over 64000000 independent realizations. The empirical means of A_n and A'_n are $2.89 \cdot 10^{-5}$ and -0.88874 and their empirical standard deviations are 0.500031 and 0.499989 . The theoretical values are $\mathbb{E}[A_n/n] = 0$, $\mathbb{E}[A'_n/(2np(1-p))] = 8/9 = 0.888\dots$ and $\sigma = 1/2$.

space $(\Omega, \mathcal{F}, \mathbb{P})$ is Λ -invariant if and only if, for all $x, y \in G$ and for all $\lambda \in \Lambda$, $Q(x + \lambda, y + \lambda) = Q(x, y)$.

Property 2.2. Let $(X_n)_{n \in \mathbb{N}}$ be a Λ -invariant Markov chain on a periodic subgraph G of E . The process $(\pi_0(X_n))_{n \in \mathbb{N}}$ is a G_0 -valued Markov chain.

Proof. Let f be any bounded Borel function $G_0 \rightarrow E$.

$$\begin{aligned} \mathbb{E}[f(\pi_0(X_{n+1}))|\mathcal{F}_n] &= \mathbb{E}\left[\sum_{\lambda \in \Lambda} \mathbf{1}_{\pi_\Lambda(X_{n+1})=\lambda} f(\pi_0(X_{n+1})) \middle| \mathcal{F}_n\right] \\ &= \sum_{\lambda \in \Lambda} \mathbb{E}[\mathbf{1}_{\pi_\Lambda(X_{n+1})=\lambda} f(\pi_0(X_{n+1}))|\mathcal{F}_n] = \sum_{\lambda \in \Lambda} (Qg_\lambda)(X_n) \end{aligned}$$

where $g_\lambda(x) = \mathbf{1}_{\pi_\Lambda(x)=\lambda} f(\pi_0(x))$ by the Markov property for (X_n) . The invariance of Q gives now:

$$\begin{aligned} (Qg_\lambda)(x) &= \sum_{y \in G} Q(x, y) g_\lambda(y) = \sum_{y_0 \in G} Q(x, y) \mathbf{1}_{\pi_\Lambda(y)=\lambda} f(\pi_0(y)) \\ &= \sum_{y_0 \in G_0} Q((\pi_\Lambda(x), \pi_0(x)), (\lambda, y_0)) f(y_0) = \sum_{y_0 \in G_0} Q((0, \pi_0(x)), (\lambda - \pi_\Lambda(x), y_0)) f(y_0) \end{aligned}$$

by Λ -invariance for Q . Summation over Λ eliminates the dependence on $\pi_\Lambda(x)$ and we thus obtain:

$$\mathbb{E}[f(\pi_0(X_{n+1}))|\mathcal{F}_n] = (Q_0 f)(\pi_0(X_n))$$

with $Q_0(x_0, y_0) = \sum_{\lambda \in \Lambda} Q((0, x_0), (\lambda, y_0))$. \square

Moreover, similar calculations show that the process $(\pi_\Lambda(X_n))_n$ knowing the process $(\pi_0(X_n))_n$ is a heterogeneous Markov chain whose rates¹ depend on the $(\pi_0(X_n))_n$.

2.2. Decomposition into pseudo-excursions. We start with a general definition of pseudo-excursions for an E -valued sequence:

Definition 2.3. Let $(x_n)_{n \in \mathbb{N}}$ be an E -valued sequence and $(T_k)_{k \in \mathbb{N}}$ be a strictly increasing sequence in \mathbb{N} such that $T_0 = 0$ and $T_{k+1} - T_k = L_k > 0$. We introduce the sequence $\tilde{\lambda}_p(x) := x_{T_p}$. Let o be an additional cemetery point added to E . The pseudo-excursions $\widetilde{\text{Exc}}^{(p)}(x)$ of the sequence $(x_n)_{n \in \mathbb{N}}$ are then defined as $E \cup \{o\}$ -valued processes through:

$$\widetilde{\text{Exc}}^{(k)}(x)_n = \begin{cases} x_{T_k+n} - \tilde{\lambda}_k(x) & \text{if } 0 \leq n \leq L_k \\ o & \text{if } n > L_k \end{cases}$$

The global trajectory $(x_n)_n$ can be recovered from the excursions by:

$$(2) \quad x_n = \tilde{\lambda}_{\kappa(n)}(x) + \widetilde{\text{Exc}}^{(\kappa(n))}(x)_{n-T_{\kappa(n)}}$$

where $\kappa(n)$ is the unique integer such that $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$.

We will now give a definition of pseudo-excursions which applies to a specific class of G -valued sequences we are interested in (with G as described in section 2.1). For this purpose, we will slightly change definition 2.3.

For $(x_n)_{n \in \mathbb{N}} \in G^{\mathbb{N}}$, we introduce the sequence of excursion times of the sequence $(\pi_0(x_n))_{n \in \mathbb{N}}$ from its original point:

¹the probability of a jump between λ and λ' between times k and $k+1$ is $Q_k(\lambda, \lambda'|\pi_0(X)) = Q((\lambda, X_k), (\lambda', X_{k+1}))/Q_0(X_k, X_{k+1})$.

$$T_0 = 0, \\ T_{k+1} = \inf \{n > T_k : \pi_0(x_n) = \pi_0(x_0)\}, \quad k \geq 0.$$

Definition 2.4. Let $(x_n)_{n \in \mathbb{N}}$ be a G -valued sequence such that $(\pi_0(x_n))_{n \in \mathbb{N}}$ is recurrent (i.e. each point of G_0 appears an infinity of times in the sequence). Set $\lambda_k(x) = \pi_\Lambda(x_{T_k})$ and $L_k = T_{k+1} - T_k$ (L_k is the duration of an excursion). Let o be an additional cemetery point added to G . The pseudo-excursions $\text{Exc}^{(k)}(x)$ of the sequence $(x_n)_{n \in \mathbb{N}}$ are defined as $G \cup \{o\}$ -valued processes through:

$$(3) \quad \text{Exc}^{(k)}(x)_n = \begin{cases} x_{T_k+n} - \lambda_k(x) & \text{if } 0 \leq n \leq L_k \\ o & \text{if } n > L_k \end{cases}$$

Although the above definition can be viewed as a particular case of definition 2.3, its interest consists in exploiting the decomposition of elements of G in $\Lambda \times G_0$ -valued couples. This enables us to translate only the Λ -valued component and thus to start each pseudo-excursion from a point $y \in G$ such that $\pi_0(y) = \pi_0(x_0)$. Moreover, as we keep here close to the classical definition of excursions, we can deal with the Markov chain $(X_n)_{n \in \mathbb{N}}$ (and not only $(\pi_0(X_n))_{n \in \mathbb{N}}$) and make computations of the Lévy area easier. In the rest of the article, we will prefer definition 2.4 when we talk of a (recurrent) G -valued sequence, and definition 2.3 will apply whenever we make a statement concerning any E -valued sequence.

One immediately checks that $\text{Exc}^{(k)}(x)_0 = \pi_0(x_0)$ and $\text{Exc}^{(k)}(x)_{L_k} = \pi_0(x_0) + \lambda_{k+1}(x) - \lambda_k(x)$. Our construction of pseudo-excursions makes the definition invariant under translation by an element of Λ : $\text{Exc}^{(k)}(\mu + x) = \text{Exc}^{(k)}(x)$ where μ is any element of Λ .

Property 2.3. Let $(X_n)_n$ be a Λ -invariant Markov chain on the periodic graph G such that the projection $(\pi_0(X_n))_n$ is irreducible. The $(G \cup \{o\})^{\mathbb{N}}$ -valued random variables $(\text{Exc}^{(k)}(X))_{k \in \mathbb{N}}$ are independent and identically distributed.

Proof. The proof relies on the repetitive use of the strong Markov property. Let $n \in \mathbb{N}^*$ and let $f_0, f_1, \dots, f_n : (G \cup \{o\})^{\mathbb{N}} \rightarrow \mathbb{R}$ be bounded measurable functions. The random times T_k are stopping times, which are finite almost surely. For $k \leq n-1$, the random variables $\mathcal{E}_k = \text{Exc}^{(k)}(X)$ are \mathcal{F}_{T_k} -measurable, and we obtain:

$$\mathbb{E}_x [f_0(\mathcal{E}_0) \dots f_{n-1}(\mathcal{E}_{n-1}) f_n(\mathcal{E}_n) | \mathcal{F}_{T_n}] = f_0(\mathcal{E}_0) \dots f_{n-1}(\mathcal{E}_{n-1}) \mathbb{E}_x [f_n(\mathcal{E}_n) | \mathcal{F}_{T_n}]$$

The strong Markov property thus yields:

$$\begin{aligned} \mathbb{E}_x [f_n(\mathcal{E}_n) | \mathcal{F}_{T_n}] &= \mathbb{E}_{X_{T_n}} [f_n(\mathcal{E}_0)] = \mathbb{E}_{(\lambda_n(x), \pi_0(x))} [f_n(\mathcal{E}_0)] \\ &= \mathbb{E}_{(\lambda_0(x), \pi_0(x))} \left[f_n(\text{Exc}^{(0)}(X + \lambda_n - \lambda_0)) \right] = \mathbb{E}_{(\lambda_0(x), \pi_0(x))} [f_n(\mathcal{E}_0)] \end{aligned}$$

a.s., where the last equality is deduced from the Λ -invariance of the process and of the pseudo-excursions. One now remarks that the last term does not depend anymore on X_{T_n} . By recursion, we obtain the final result:

$$\mathbb{E}_x [f_0(\mathcal{E}_0) \dots f_{n-1}(\mathcal{E}_{n-1}) f_n(\mathcal{E}_n)] = \mathbb{E}_x [f_0(\mathcal{E}_0)] \mathbb{E}_x [f_1(\mathcal{E}_0)] \mathbb{E}_x [f_n(\mathcal{E}_0)]$$

□

Corollary 2.1. *The random variables $(\lambda_{k+1}(X) - \lambda_k(X))_{k \in \mathbb{N}}$ are also i.i.d.*

Proof. This follows directly from $\lambda_{k+1}(X) - \lambda_k(X) = \text{Exc}^{(k)}(X)_{L_k} - \text{Exc}^{(k)}(X)_0$ and the independence of the pseudo-excursions. \square

Remark: If G is a periodic graph and $M \in GL_n(\mathbb{R})$, then $MG = \{Mx; x \in G\}$ is again a periodic graph (with possibly degenerate vertices). If $(X_n)_n$ is a Λ -invariant Markov chain on G , then $(MX_n)_n$ is again a $M\Lambda$ -invariant Markov chain on MG . We assume all through the paper that $\mathbb{R}^n = \text{span } \Lambda$; if this is not the case, we embed the graph G in the smaller space $\text{span } \Lambda$ isomorphic to some \mathbb{R}^n . Let C be the covariance matrix of the increment $\lambda_1(X) - \lambda_0(X)$, we always assume that $C = M^* I_n M$. If $\text{rank } C < n$, we again embed our Markov chain in a smaller graph in a smaller space such that $\text{rank } C = n$. Then, up to reduction to a smaller space and up to an invertible linear transformation of the graph, we may always assume that $C = I_n$. In particular, under the $C = I_n$ hypothesis, the Donsker embedding of the random walk $(\lambda_n(x))_n$ converges to a standard Brownian motion on \mathbb{R}^n .

Example of section 1.3. The Markov chain (Z'_n) fits into this framework with $G = \mathbb{Z}^2$. The non-zero elements of the transition matrix Q are represented in figure 1. The matrix Q is Λ -invariant with $\Lambda = (2\mathbb{Z})^2$. The set $G_0 = \{(0,0), (1,0), (0,1), (1,1)\}$ may be identified to $\mathbb{Z}/4\mathbb{Z}$, and so the Markov process $(\pi_0(Z'_n))_n$ is actually deterministic and corresponds to the shift $x \mapsto x + 1$ as in figure 1.

2.3. Area process and rough paths.

Definition 2.5 (area sequence). *Let E be a finite-dimensional vector space. Let $(e_i)_{1 \leq i \leq d}$ be a basis of E . We write $x^{(i)}$ for the i -th coordinate of a vector $x \in E$ w.r.t. the basis $(e_i)_{1 \leq i \leq d}$. For any E -valued sequence $(x_n)_{n \in \mathbb{N}}$, we introduce the sequence of antisymmetric $d \times d$ matrices $(A_n(x))_{n \in \mathbb{N}}$ defined by $A_0(x) = A_1(x) = 0$ and, for any $n \geq 2$,*

$$(4) \quad A_n^{ij}(x) = \sum_{1 \leq k < l \leq n} \left((\Delta x^{(i)})_k (\Delta x^{(j)})_l - (\Delta x^{(j)})_k (\Delta x^{(i)})_l \right)$$

with $(\Delta u)_k = u_k - u_{k-1}$ for any sequence $(u_n)_n$.

Property 2.4 (decomposition of an area sequence along excursions). *Let $(x_n)_{n \in \mathbb{N}}$ and $(T_k)_{k \in \mathbb{N}}$ be as in definition 2.3. Then the following decomposition holds:*

$$(5) \quad A_{T_n}^{ij}(x) = \sum_{p=0}^{n-1} A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(x)) + A_n^{ij}(\tilde{\lambda}(x))$$

In the particular case when $(x_n)_{n \in \mathbb{N}}$ and $(T_k)_{k \in \mathbb{N}}$ are as in definition 2.4, we have:

$$(6) \quad A_{T_n}^{ij}(x) = \sum_{p=0}^{n-1} A_{L_p}^{ij}(\text{Exc}^{(p)}(x)) + A_n^{ij}(\lambda(x))$$

Proof. By definition, the l.h.s. uses a double sum over $1 \leq k < l \leq T_n$. We split the interval $\{1, 2, \dots, T_n\}$ into $J_p = \{T_p + 1, \dots, T_{p+1}\}$ for $p = 0, \dots, n-1$ and we

classify the indices k and l : either they are in the same subset J_p or they belong respectively to J_{p_1} and J_{p_2} with $p_1 < p_2$.

In the first case, the sum over $T_r + 1 \leq k < l \leq T_{r+1}$ gives the area of the r -th excursion $A_{L_r}^{ij}(\text{Exc}^{(r)}(x))$.

In the second case, the sum over $k \in J_{p_1}$ and $l \in J_{p_2}$ factorizes into two sums, evaluated as telescopic sums respectively to $\lambda_{p_i+1}(x) - \lambda_{p_i}(x)$ for $i = 1, 2$. The remaining sum over $0 \leq p_1 < p_2 \leq n - 1$ gives the (signed) area of $(\lambda_k(x))_{k \in \mathbb{N}}$ between 0 and n . \square

We need a last lemma, easy to prove, from linear algebra, about the transformation of the area under a linear transformation M of the space E

Lemma 2.1 (covariance of the area). *Let $(x_n)_n$ be an E -valued sequence and $M \in GL_n(E)$. The area process $(A_n^{ij}(x))_n$ of the sequence $(Mx_n)_n$ in E is given by:*

$$(7) \quad A_n^{ij}(Mx) = \sum_{1 \leq k, l \leq n} M_{ik} M_{jl} A_n^{kl}(x)$$

2.4. Embeddings of the discrete processes in continuous time. For any $G_2(E)$ -valued sequence $(x_n, a_n)_{n \in \mathbb{N}}$, we define its embedding in continuous time by linear and geodesic interpolation:

$$\iota^{(N)}(x_\bullet, a_\bullet)_t = \left(\frac{x_{\lfloor Nt \rfloor} + (Nt - \lfloor Nt \rfloor)(x_{\lfloor Nt \rfloor + 1} - x_{\lfloor Nt \rfloor})}{\sqrt{N}}, \right. \\ \left. \frac{a_{\lfloor Nt \rfloor} + (Nt - \lfloor Nt \rfloor)(a_{\lfloor Nt \rfloor + 1} - a_{\lfloor Nt \rfloor})}{N} \right)$$

as it is usual for the formulation of Donsker theorem for random walks.

3. CONVERGENCE OF THE DISCRETE PROCESS TO THE ANOMALOUS BROWNIAN MOTION

3.1. Notations. All the random variables used below are defined in the previous section and all the topologies are described in section 5.1. We denote by $|\cdot|$ the absolute value on \mathbb{R} , by $|\cdot|_E$ the euclidean norm on the finite-dimensional vector space E and by $|\cdot|_{E \otimes E}$ the induced matrix norm on $E \otimes E$: $|A|_{E \otimes E} = \sup_{|x|_E=1} |Ax|_E$. We also use

the norm $\|\cdot\|$ on $G_2(E)$ and the associated distance $\mathbf{d}(\cdot, \cdot)$ (defined in section 5.2.2). We set, for $u, v \in G_2(E)^l$, $\mathbf{d}_l((u^1, \dots, u^l), (v^1, \dots, v^l)) = \sum_{i=1}^l \mathbf{d}(u^i, v^i)$: \mathbf{d}_l is a distance on $G_2(E)^l$.

For any $n \in \mathbb{N}$, we define $\kappa(n)$ as the unique integer such that $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$, where the T_n 's are as in definitions 2.3 or 2.4 (as has already been done in section 2.2).

We denote by δ_ϵ the standard homogeneous dilatation on $G_2(E)$, i.e. $\delta_\epsilon(x, a) = (\epsilon x, \epsilon^2 a)$.

3.2. Statement of the main theorem. We may now proceed to studying the scaling limit of the process $(X_n, A_n(X))_n$ in the rough path topology.

Theorem 3.1. *Let G be a Λ -periodic graph in a finite dimensional vector space E . Let $(X_n)_n$ be a G -valued Λ -invariant Markov chain on G with bounded increments, (i.e. there exists $R > 0$ such that $|X_{n+1} - X_n|_E \leq R$ a.s.) and such that $(\pi_0(X_n))_{n \in \mathbb{N}}$ is irreducible (since G_0 is finite, this implies in particular that for any initial law μ on G_0 , T_1 has finite moments of all orders).*

Let $v = \mathbb{E}[T_1]^{-1} \mathbb{E}[X_{T_1}] \in E$ and $\beta = \mathbb{E}[T_1] \in \mathbb{R}_+^*$. Let $(\tilde{X}_n)_{n \in \mathbb{N}}$ be the E -valued process defined by $\tilde{X}_n = X_n - nv$. Up to a dimensional reduction and a linear transformation of the graph G , the covariance matrix of $\lambda_1(X) - T_1 v$ may always be assumed to be CI_n with $C > 0$.

For any $\tau > 0$, the sequence of processes $\left(\delta_{\sqrt{C^{-1}\beta}} \iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t \right)_{0 \leq t \leq \tau}$ converges in distribution, as $N \rightarrow \infty$, to the process $(B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau}$ in the topology of $C^{0,\alpha}\text{-Holder}([0, \tau], G_2(E))$ for $\alpha < 1/2$, where B is a standard Brownian motion on E , \mathcal{A} its Lévy area as defined by classical stochastic calculus and Γ a constant antisymmetric matrix, the area anomaly, given by (9).

3.3. Consequences of theorem 3.1. The hypothesis of theorem 3.1 are satisfied in many models coming from statistical mechanics where jumps in space are often local. Up to our knowledge, the area anomaly is a new feature never described in any model, even if the examples that we present look natural. One may wonder whether this area anomaly is relevant. We now explain why it is the case.

The general philosophy beyond renormalization and large scale limits of discrete models is to build continuous models such that they are large scale limits of various discrete models and such that it is possible to compute directly with them.

Donsker theorem for example states that, after renormalization, it is more convenient to work with Brownian motion instead of random walks with possibly complicated jump laws. Our theorem states that this is indeed the case for 1-dimensional-space valued Markov chains, for which the concept of Lévy area is absent but this is *not* anymore the case for d -dimensional-space-valued Markov chains, for which the Lévy area may have an anomaly Γ .

Phrased in a more provocative way, we could write that *a two-dimensional standard Brownian motion may not be the same as two independent one-dimensional Brownian motion* as soon as one wishes to use it to drive a stochastic differential equation. The difference lies in the area anomaly Γ which is irrelevant at the level of the positions $(B_{t_k})_k$ but is irrelevant in non-linear SDEs.

As soon as several Brownian motions emerge in the description of the limit of discrete processes, the consequence of the previous theorem is that one *needs in general to wonder* about the presence of area anomalies between components before writing down any stochastic integration.

Hopefully in many cases, it is easy to prove without any calculation that the area anomaly is zero. If the discrete model is reversible, its limit has to be also reversible and thus the area anomaly has to be zero. However, for irreversible Markov chains, especially useful in non-equilibrium statistical mechanics, one should expect in general a non-zero anomaly.

The full study of the area anomaly Γ and its generalization to more general processes will be present in a next paper in preparation.

3.4. Proof of the main theorem. Since $\tilde{\lambda}_k(\tilde{X}) = \lambda_k(X) - T_k v$ (with $\lambda_k(X)$ and $\tilde{\lambda}_k(\tilde{X})$ as in definitions 2.4 and 2.3 respectively), the process $(\tilde{\lambda}_k(\tilde{X}))_k$ is an E -valued centred random walk (not Λ -valued because of the correction) and, by assumption on the normalization of the graph G , each (E -valued) increment has a covariance matrix equal to CI_n .

The main idea of the proof of theorem 3.1 is to use the theory of pseudo-excursions from section 2.2 and the decomposition from property 2.4 in order to extract convergence to the standard Brownian rough path through the process

$(\tilde{\lambda}_n(\tilde{X}))_{n \in \mathbb{N}}$ using [4], convergence to the area anomaly through the independence of pseudo-excursions, and tightness from additional results on pseudo-excursions. Consequently, the proof of theorem 3.1 is divided into 4 steps:

- lemma 3.1: convergence of the centred discrete process $(\iota^{(N)}(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{\lambda}(\tilde{X})))_t)_{0 \leq t \leq \tau}$
- lemma 3.2: convergence of the extracted process $(\iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X})))_t)_{0 \leq t \leq \tau}$ and emergence of the area anomaly (drift) Γ
- lemma 3.3: convergence of finite-dimensional marginals of the full process $(\iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X})))_t)_{0 \leq t \leq \tau}$
- lemma 3.4: tightness of the sequence $(\iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X})))_{N \in \mathbb{N}}$

Lemma 3.1. *The process $(\iota^{(N)}(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{\lambda}(\tilde{X})))_t)_{0 \leq t \leq \tau}$ converges in distribution to the Lévy lift on $G_2(E)$ of a Brownian motion $(B_t)_{t \geq 0}$:*

$$(\delta_{\sqrt{C^{-1}}} \iota^{(N)}(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{\lambda}(\tilde{X})))_t)_{0 \leq t \leq \tau} \xrightarrow[N \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t)_{0 \leq t \leq \tau}$$

in the topology of $C^{0, \alpha\text{-Holder}}([0, \tau], G_2(E))$ for $\alpha < 1/2$.

Proof. This is a direct consequence of the Donsker-type theorem for a sequence of i.i.d. centred $G_2(\mathbb{R}^2)$ -valued random variables from [4]. In this article, the authors use a central limit theorem for centred i.i.d. variables on a nilpotent Lie group in order to prove the convergence of finite-dimensional distributions, and Kolmogorov's criterion to prove the tightness of the sequence. \square

Lemma 3.2. *The sequence of processes $(\delta_{\sqrt{C^{-1}}} \iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X})))_t)_{0 \leq t \leq \tau}$ converges in distribution to $(B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau}$ in the topology of $C^{0, \alpha\text{-Holder}}([0, \tau], G_2(E))$ for $\alpha < 1/2$, with Γ given by (9).*

This is the part of the proof where the area anomaly Γ first appears. We will see that, between 3.1 and 3.2, nothing changes on the first level of the new sequence, since the embedding is obtained by linear interpolation and therefore does not keep track of the trajectory between T_n and T_{n+1} . Simultaneously, a complementary term appears on the second level, in the expression of the stochastic area. This is due to the fact that whereas the specific trajectory of an excursion is not memorized, its area is registered in the continuous embedding.

Proof. We have trivially by definition 2.3:

$$\tilde{X}_{T_n} = \tilde{\lambda}_n(\tilde{X})$$

Moreover, property 2.4 applied to \tilde{X} gives:

$$(8) \quad A_{T_n}^{ij}(\tilde{X}) = A_n^{ij}(\tilde{\lambda}(\tilde{X})) + \sum_{p=0}^{n-1} A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X}))$$

Each term $A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X}))$ represents exactly the area of the $(p+1)$ -th excursion and the total sum is the complementary second-level term mentioned above.

Let us decompose using (4):

$$\begin{aligned}
A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X})) &= \sum_{1 \leq k < l \leq L_p} \left((\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(i)})_k (\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(j)})_l \right. \\
&\quad \left. - (\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(j)})_k (\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(i)})_l \right) \\
&= A_{L_p}^{ij}(\text{Exc}^{(p)}(X)) \\
&\quad + \left(\sum_{1 \leq k < l \leq L_p} ((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1})) \right)^{(i)} v^{(j)} \\
&\quad - v^{(i)} \left(\sum_{1 \leq k < l \leq L_p} ((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1})) \right)^{(j)}
\end{aligned}$$

We set

$$\begin{aligned}
\mathbf{Corr}_p^{ij}(X) &= \left(\sum_{1 \leq k < l \leq L_p} ((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1})) \right)^{(i)} v^{(j)} \\
&\quad - v^{(i)} \left(\sum_{1 \leq k < l \leq L_p} ((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1})) \right)^{(j)}
\end{aligned}$$

and we call this term the *area drift correction*. Since we have supposed that the increments of X are bounded by a certain $R > 0$, we deduce that

$$|\mathbf{Corr}_p^{ij}(X)| \leq K_{v,R} L_p^2$$

where K is a constant depending on v and R . Likewise, we obtain

$$|A_{L_p}^{ij}(\text{Exc}^{(p)}(X))| \leq K'_R L_p^2$$

where K' is a constant depending on R . We can thus conclude that all the $A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X}))$ are integrable. Moreover, these variables are i.i.d., since $A_{L_p}^{ij}(\text{Exc}^{(p)}(X))$ and $\mathbf{Corr}_p^{ij}(X)$ depend only on $\text{Exc}^{(p)}(X)$. Thus, by the law of large numbers the following convergence holds:

$$\frac{1}{n} \sum_{p=0}^{n-1} A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X})) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathbb{E} \left[A_{L_0}^{ij}(\widetilde{\text{Exc}}^{(0)}(\tilde{X})) \right] = \mathbb{E} \left[A_{L_0}^{ij}(\text{Exc}^{(0)}(X)) \right] + \mathbb{E} \left[\mathbf{Corr}_0^{ij}(X) \right]$$

Finally, using the result from step 1 and Slutsky's theorem, we can conclude that

$$\left(\delta_{\sqrt{C^{-1}}} \iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X}))_t \right)_{0 \leq t \leq \tau} \xrightarrow[N \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau}$$

where the coefficients of the $d \times d$ (with $d = \dim(E)$) matrix Γ are given by

$$(9) \quad \Gamma^{ij} = C^{-1} (\mathbb{E} [A_{L_0}^{ij}(\text{Exc}^{(0)}(X))] + \mathbb{E} [\mathbf{Corr}_0^{ij}(X)])$$

The matrix Γ is the announced area anomaly. It is immediate from definition 4 that $A_{L_0}^{ij}(\widetilde{\text{Exc}}^{(0)}(\tilde{X})) = -A_{L_0}^{ji}(\widetilde{\text{Exc}}^{(0)}(\tilde{X}))$, which implies that Γ is antisymmetric. \square

Lemma 3.3. *For any $t_1 < t_2 < \dots < t_k \in \mathbb{R}_+$, we have*

$$\begin{aligned} & \left(\delta_{\sqrt{C^{-1}\beta}} \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{t_1}, \dots, \delta_{\sqrt{C^{-1}\beta}} \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{t_k} \right) \\ & \xrightarrow[n \rightarrow \infty]{(d)} ((B_{t_1}, \mathcal{A}_{t_1} + t_1 \Gamma), \dots, (B_{t_k}, \mathcal{A}_{t_k} + t_k \Gamma)) \end{aligned}$$

In this lemma, we pass from the embeddings of an extracted sequence $(\iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X}))_t)_{0 \leq t \leq \tau}$ to the embeddings of the full sequence $(\iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t)_{0 \leq t \leq \tau}$. We show that in the term $\delta_{\sqrt{n^{-1}}}(\tilde{X}_{[nt]}, A_{[nt]}(\tilde{X}))$ the only part that counts at the limit is the one given by the excursions up to time n , i.e. $\delta_{\sqrt{n^{-1}}}(\tilde{X}_{\kappa([nt])}, A_{\kappa([nt])}(\tilde{X}))$. At the same time, the constant β appears in the renormalization, since we have to take into consideration the approximate length of an excursion up to time $n \frac{n}{\kappa(n)}$, and β is precisely the a.s. limit of this sequence.

Proof. Set $\tilde{\mathbb{X}}_n = (\tilde{X}_n, A_n(\tilde{X}))$. With the upper bound from proposition 5.1 (since it doesn't play an important role here, we suppose that $\nu = 1$), for $t \in [0, \tau]$, we get the inequality:

$$\mathbf{d}(\delta_{\sqrt{n^{-1}}} \tilde{\mathbb{X}}_{T_{\kappa([nt])}}, \delta_{\sqrt{n^{-1}}} \tilde{\mathbb{X}}_{[nt]}) \leq \frac{1}{\sqrt{n}} |\tilde{X}_{[nt]} - \tilde{X}_{T_{\kappa([nt])}}|_E + \frac{1}{\sqrt{n}} |A_{T_{\kappa([nt])}, [nt]}(\tilde{X})|_{E \otimes E}^{\frac{1}{2}}$$

We are going to use this decomposition in order to prove that $\mathbf{d}(\delta_{\sqrt{n^{-1}}} \tilde{\mathbb{X}}_{T_{\kappa([nt])}}, \delta_{\sqrt{n^{-1}}} \tilde{\mathbb{X}}_{[nt]})$ converges in probability to 0.

We set $\tilde{R} = R + |v|_E$. First, it is easy to see that if $k' \leq k < k''$, by triangular inequality we have a.s.

$$(10) \quad |\tilde{X}_{k-k'}|_E \leq \sum_{l=1}^{k-k'} |\tilde{X}_l - \tilde{X}_{l-1}|_E \leq \tilde{R}(k - k') \leq \tilde{R}(k'' - k')$$

Next, since $A_n(\tilde{X})$ is a $d \times d$ matrix (with $d = \dim(E)$), we have

$$(11) \quad |A_{k-k'}(\tilde{X})|_{E \otimes E} \leq \sum_{i=1}^d \max_{j=1, \dots, d} |A_{k-k'}^{ij}(\tilde{X})| \leq d\tilde{R}^2(k - k')^2 \leq d\tilde{R}^2(k'' - k')^2$$

Further on, by strong Markov property, for $\epsilon > 0$, the Chebyshev's inequality, together with (10), implies, for the first term,

$$\begin{aligned} \mathbb{P}\left(\frac{1}{\sqrt{n}} |\tilde{X}_{[nt]} - \tilde{X}_{T_{\kappa([nt])}}|_E > \epsilon\right) & \leq \frac{\mathbb{E}\left[\left|\tilde{X}_{[nt]} - \tilde{X}_{T_{\kappa([nt])}}\right|_E^2\right]}{n\epsilon^2} \leq \frac{\tilde{R}^2 \mathbb{E}[(T_{\kappa([nt])} + 1 - T_{\kappa([nt])})^2]}{n\epsilon^2} \\ & = \frac{\tilde{R}^2 \mathbb{E}[T_1^2]}{n\epsilon^2} \end{aligned}$$

and, for the second term, together with (11),

$$\mathbb{P}\left(\frac{1}{\sqrt{n}}\left|A_{T_{\kappa(\lfloor nt \rfloor)}, \lfloor nt \rfloor}(\tilde{X})\right|_{E \otimes E}^{\frac{1}{2}} > \epsilon\right) \leq \frac{\mathbb{E}\left[\left|A_{T_{\kappa(\lfloor nt \rfloor)}, \lfloor nt \rfloor}(\tilde{X})\right|_{E \otimes E}\right]}{n\epsilon^2} \leq \frac{d\tilde{R}^2\mathbb{E}[T_1^2]}{n\epsilon^2}$$

Hence, taking into consideration the fact that $\mathbb{E}[T_1^2] < \infty$, we obtain

$$\begin{aligned} \mathbb{P}\left(\mathbf{d}\left(\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt \rfloor)}}, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt \rfloor}\right) > \epsilon\right) &\leq \mathbb{P}\left(\frac{1}{\sqrt{n}}\left|\tilde{X}_{\lfloor nt \rfloor} - \tilde{X}_{T_{\kappa(\lfloor nt \rfloor)}}\right|_E > \frac{\epsilon}{2}\right) \\ &+ \mathbb{P}\left(\frac{1}{\sqrt{n}}\left|A_{T_{\kappa(\lfloor nt \rfloor)}, \lfloor nt \rfloor}(\tilde{X})\right|_{E \otimes E}^{\frac{1}{2}} > \frac{\epsilon}{2}\right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

As $\kappa(n)$ is the number of excursions up to a time n of $(\pi_0(X_n))_n$, the ergodic theory tells us that $\frac{\kappa(n)}{n} \xrightarrow{n \rightarrow \infty} \frac{1}{\beta}$ a.s. Consequently, the above convergence in probability combined with the result from step 2 implies

$$\delta_{\sqrt{n-1}C^{-1}\beta}\tilde{\mathbb{X}}_{\lfloor nt \rfloor} \xrightarrow[n \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t + t\Gamma)$$

What is now left to do is pass from $\tilde{\mathbb{X}}_{\lfloor nt \rfloor}$ to $\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t$, and in order to do that we have to study the convergence of $\frac{|\tilde{X}_{\lfloor nt \rfloor+1} - \tilde{X}_{\lfloor nt \rfloor}|_E}{\sqrt{n}}$ and $\frac{|A_{\lfloor nt \rfloor+1}(\tilde{X}) - A_{\lfloor nt \rfloor}(\tilde{X})|_{E \otimes E}}{n}$.

We start with $\frac{|\tilde{X}_{\lfloor nt \rfloor+1} - \tilde{X}_{\lfloor nt \rfloor}|_E}{\sqrt{n}} \leq \frac{\tilde{R}}{\sqrt{n}} \xrightarrow{n \rightarrow \infty} 0$ a.s. Similarly, by the decomposition 5 from lemma 5.1 and using the fact that $\frac{1}{2}|ab - cd| \leq a^2 + b^2 + c^2 + d^2$,

$$\begin{aligned} \frac{|A_{\lfloor nt \rfloor+1}(\tilde{X}) - A_{\lfloor nt \rfloor}(\tilde{X})|_{E \otimes E}}{n} &\leq \frac{\tilde{R}^2 + |\tilde{X}_{\lfloor nt \rfloor} - \tilde{X}_{T_{\kappa(\lfloor nt \rfloor)}}|_E^2}{n} \\ &\leq \frac{\tilde{R}^2((T_{\kappa(\lfloor nt \rfloor)+1} - T_{\kappa(\lfloor nt \rfloor)})^2 + 1)}{n} \xrightarrow[n \rightarrow \infty]{} 0 \text{ in probability} \end{aligned}$$

We conclude by Slutsky's theorem that

$$\delta_{\sqrt{C^{-1}\beta}}\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t \xrightarrow[n \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t + t\Gamma)$$

It is now easy to pass to the multivariate case. Choose $t_1 < t_2 < \dots < t_l \in \mathbb{R}_+$. Then we have immediately

$$\begin{aligned} \mathbb{P}\left(\mathbf{d}_l\left((\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt_1 \rfloor)}}, \dots, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt_l \rfloor)}}), (\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt_1 \rfloor}, \dots, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt_l \rfloor})\right) > \epsilon\right) \\ \leq \sum_{i=1}^l \mathbb{P}\left(\mathbf{d}\left(\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt_i \rfloor}, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt_i \rfloor)}}\right) > \frac{\epsilon}{l}\right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

Applying once again the result from step 2, we obtain

$$\left(\delta_{\sqrt{n-1}C^{-1}\beta}\tilde{\mathbb{X}}_{t_1}, \dots, \delta_{\sqrt{n-1}C^{-1}\beta}\tilde{\mathbb{X}}_{t_k}\right) \xrightarrow[n \rightarrow \infty]{(d)} ((B_{t_1}, \mathcal{A}_{t_1} + t_1\Gamma), \dots, (B_{t_k}, \mathcal{A}_{t_k} + t_k\Gamma))$$

and we conclude by applying Slutsky's theorem as in the univariate case. \square

Lemma 3.4. *The sequence $(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X})))_{n \geq 0}$ is tight in α -Hölder topology for $\alpha < 1/2$.*

Proof. As in [4], we apply here the Kolmogorov's criterion. In order to do so, it will be enough to prove that, for $\tau > 0$ fixed, for any $p > 1$ there exists a positive constant c_p such that, for all $0 \leq s < t \leq \tau$,

$$\sup_n \mathbb{E} \left[\mathbf{d} \left(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t, \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_s \right)^{4p} \right] \leq c_p |t - s|^{2p}$$

since $\frac{2p-1}{4p} \xrightarrow{p \rightarrow \infty} \frac{1}{2}$.

Choose $a > 0$. By proposition 5.1 and applying to $(X_n)_{n \in \mathbb{N}}$ the Markov property, we get:

$$\begin{aligned} \mathbb{E} \left[\mathbf{d} \left(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t, \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_s \right)^a \right] &= \mathbb{E} \left[\left| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{s,t} \right|^a \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\left| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{s,t} \right|^a \middle| \tilde{X}_{[ns]} \right] \right] = \mathbb{E} \left[\mathbb{E}_{\tilde{X}_{[ns]}} \left[\left| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{t-s} \right|^a \right] \right] \\ &= \mathbb{E} \left[\left| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{t-s} \right|^a \right] \end{aligned}$$

Since $(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t)_{0 \leq t \leq \tau}$ is constructed by linear connections between the points $\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{\frac{k}{n}}$ for $k = 0, \dots, \lfloor n\tau \rfloor$, the properties of geodesic interpolation imply that it is sufficient to prove that

$$\frac{1}{n^{2p}} \mathbb{E} \left[\left| \tilde{\mathbb{X}}_k \right|^{4p} \right] \leq c_p \left(\frac{k}{n} \right)^{2p}$$

for $k = 0, \dots, \lfloor n\tau \rfloor$, uniformly over $n \geq 1$. As in [4], this follows immediately if we prove, for all $p > 1$,

$$\mathbb{E} \left[\left| \tilde{\mathbb{X}}_n \right|^{4p} \right] = O(n^{2p})$$

Here, Chen's relation (formula (14)) gives

$$\tilde{\mathbb{X}}_n = \tilde{\mathbb{X}}_{T_{\kappa(n)}} \otimes_2 \tilde{\mathbb{X}}_{T_{\kappa(n)}, n}$$

where \otimes_2 is the product on $G_2(E)$ from section 5.2.1 (it can also be interpreted as a path concatenation operator). As mentioned in section 5.2.2, the norm $\|\cdot\|$ is sub-additive. Using strong Markov property and the inequality: $\forall a, b \geq 0, (a+b)^p \leq 2^p(a^p + b^p)$ (\star), we arrive to an initial upper bound:

$$\mathbb{E} \left[\left| \tilde{\mathbb{X}}_n \right|^{4p} \right] \leq 2^{4p} \left(\mathbb{E} \left[\left| \tilde{\mathbb{X}}_{T_{\kappa(n)}} \right|^{4p} \right] + \mathbb{E} \left[\left| \tilde{\mathbb{X}}_{T_{\kappa(n)}, n} \right|^{4p} \right] \right)$$

On one hand, as $\kappa(n) \leq n$ a.s., we have

$$\mathbb{E} \left[\left| \tilde{\mathbb{X}}_{T_{\kappa(n)}} \right|^{4p} \right] \leq \max_{l=1, \dots, n} \mathbb{E} \left[\left| \tilde{\mathbb{X}}_{T_l} \right|^{4p} \right] = O(n^{2p})$$

since \tilde{X}_{T_l} is a product of l centred i.i.d. variables ($\tilde{X}_{T_l} = \tilde{X}_{T_1} \otimes_2 \tilde{X}_{T_1, T_2} \otimes_2 \dots \otimes_2 \tilde{X}_{T_{l-1}, T_l}$), and therefore $\mathbb{E} [|\tilde{X}_{T_l}|^{4p}] = O(l^{2p})$ as it was proved in [4].

On the other hand, proposition 5.1 (with, once again, the convention $\nu = 1$), and the inequality (\star) , together with the upper bounds from (10) and (11), give

$$\mathbb{E} [|\tilde{X}_{T_{\kappa(n), n}}|^{4p}] \leq 2^{4p} \tilde{R}^{4p} (d^{2p} + 1) \mathbb{E} [(T_{\kappa(n)+1} - T_{\kappa(n)})^{4p}] = 2^{4p} \tilde{R}^{4p} (d^{2p} + 1) \mathbb{E} [T_1^{4p}]$$

We therefore obtain

$$\mathbb{E} [|\tilde{X}_n|^{4p}] \leq 2^{4p} \left(O(n^{2p}) + 2^{4p} \tilde{R}^{4p} (d^{2p} + 1) \mathbb{E} [T_1^{4p}] \right) = O(n^{2p})$$

which achieves the proof. \square

4. EXAMPLES

This section presents additional examples of processes in dimensions two and three which exhibit a non-zero area anomaly Γ .

4.1. The diamond model in dimension two. The example presented in the introduction can be easily extended to other regular lattices (hexagonal and triangular) in dimension 2 by using other roots of unity. We present now another less trivial type of regular lattice which exhibits a non-zero area anomaly and which is frequently used in statistical mechanics.

The graph G is defined by

$$(12) \quad G = \{(\pm L_1 + L'_1 k_1, L'_2 k_2), (L'_1 k_1, \pm L_2 + L'_2 k_2); (k_1, k_2) \in \mathbb{Z}^2\} \subset \mathbb{R}^2$$

and is represented in figure 3. We set $L'_k = D_k + 2L_k$ with $D_k \geq 0$. The fundamental domain G_0 , the "diamond", contains the four elements $a_1 = (L_1, 0)$, $a_2 = (0, L_2)$, $a_3 = (-L_1, 0)$ and $a_4 = (0, -L_2)$ and the lattice is $\Lambda = L'_1 \mathbb{Z} \times L'_2 \mathbb{Z}$. The only non-zero transition probabilities are given, up to translation by $L'_1 \mathbb{Z} \times L'_2 \mathbb{Z}$, by:

$$\begin{aligned} Q((L_1, 0), (0, L_2)) &= Q((-L_1, 0), (0, -L_2)) = p \\ Q((0, L_2), (-L_1, 0)) &= Q((0, -L_2), (L_1, 0)) = p \\ Q((L_1, 0), (L_1 + D_1, 0)) &= Q((-L_1, 0), (-L_1 - D_1, 0)) = 1 - p \\ Q((0, L_2), (0, L_2 + D_2)) &= Q((0, -L_2), (0, -L_2 - D_2)) = 1 - p \end{aligned}$$

When p increases to 1, the chain tends to stay in the diamond for a longer time and develop its area because of the spatial extension of the diamond. A numerically-generated typical trajectory is presented in figure 4 In the basis given by the a_k s, the transition matrix of the projection Q_0 is:

$$Q_0 = \begin{pmatrix} 0 & p & (1-p) & 0 \\ 0 & 0 & p & (1-p) \\ (1-p) & 0 & 0 & p \\ p & (1-p) & 0 & 0 \end{pmatrix}$$

The symmetries of the process imply a zero drift of the chain X_n and a uniform invariant measure on G_0 . However, the area anomaly is non zero.

Simulations are made for $p = 3/4$ run over 64.10^6 simulations and the process is observed at time $n = 40000$. We obtain the following values for both coordinates $X_n^{(i)} / \sqrt{n}$:

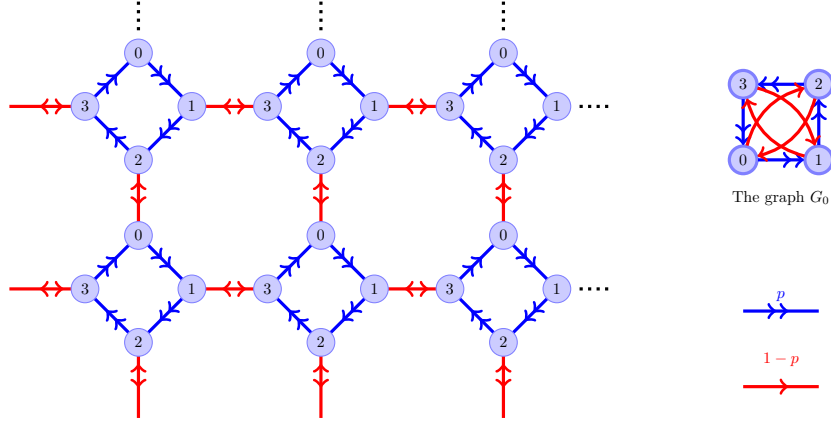
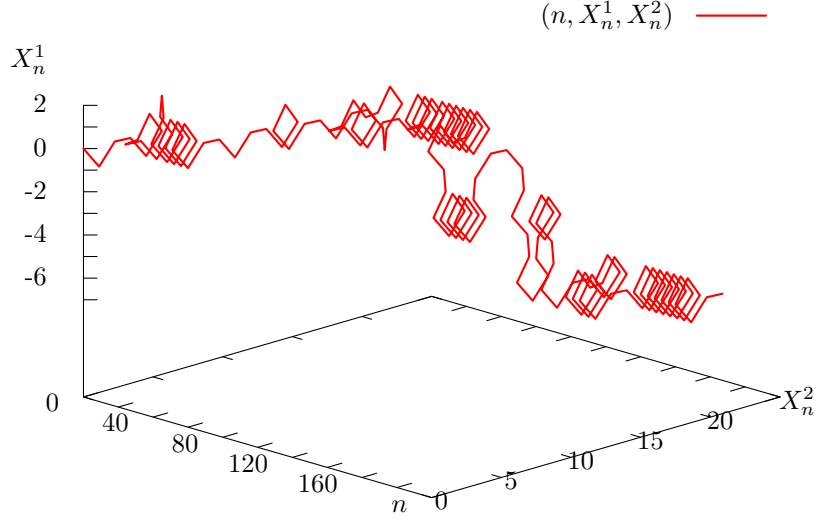


FIGURE 3. Periodic diamond graph and the loop-biased dynamics.

FIGURE 4. A typical trajectory on the diamond graph with $(X_n)_{0 \leq n \leq 200}$ with $p = 0.85$. A significative amount of time is spent turning around the diamonds.

- the empirical means are 0.0029 and -0.0014 ,
- the empirical standard deviations σ are both 0.891,
- the empirical third moments are 0.0066 and -0.0032 ,
- the kurtosis are both 3.000.

The empirical values for the area $A_n^{12}/(\sigma^2 n)$ normalized by the empirical standard deviations of the coordinates are:

- empirical mean 0.3472 (the area anomaly),
- empirical standard deviation is 0.500,
- empirical third cumulant $3.6 \cdot 10^{-5}$,
- empirical fourth cumulant 0.12505.

Origin in G	Proj. on G_0	$+e_1$	$-e_1$	$+e_2$	$-e_2$	$+e_3$	$-e_3$
$(2k, 2l, 2m)$	$(0, 0, 0)$	$u/2$	$v/2$	$u/2$	$v/2$	0	0
$(2k+1, 2l, 2m)$	$(1, 0, 0)$	0	0	$u/2$	$v/2$	$u/2$	$v/2$
$(2k, 2l+1, 2m)$	$(0, 1, 0)$	$u/3$	$v/3$	$v/3$	$u/3$	$u/3$	$v/3$
$(2k+1, 2l+1, 2m)$	$(1, 1, 0)$	$v/2$	$u/2$	0	0	$u/2$	$v/2$
$(2k, 2l, 2m+1)$	$(0, 0, 1)$	$u/2$	$v/2$	0	0	$v/2$	$u/2$
$(2k+1, 2l, 2m+1)$	$(1, 0, 1)$	$v/3$	$v/3$	$u/3$	$v/3$	$v/3$	$u/3$
$(2k, 2l+1, 2m+1)$	$(0, 1, 1)$	0	0	$v/2$	$u/2$	$v/2$	$u/2$
$(2k+1, 2l+1, 2m+1)$	$(1, 1, 1)$	$v/2$	$u/2$	$v/2$	$u/2$	0	0

FIGURE 5. Parameter of the dynamics of the cubic model used for the numerical results. The numerical simulations are made for $u = 9/10$ and $v = 1 - u = 1/10$ in order to bias the Markov chain to stay in a cube, so that it can develop a non-zero area anomaly Γ .

All the cumulants corresponds to a normal law for the coordinates and a Lévy drifted area for the area process. Full histograms, simulations and precise data are available on the webpage of the authors.

4.2. A three dimensional model with a non-trivial area anomaly. We extend the model presented in the introduction to dimension three. This extension is interesting for two main reasons: no particular role is played by the roots of unity as in the introduction and we may choose arbitrary jump rates; moreover, the area anomaly is now an antisymmetric three-by-three matrix which can be arbitrary.

The graph G is \mathbb{Z}^3 , the lattice Λ is $(2\mathbb{Z})^3$ and the fundamental domain is thus $G_0 = (\mathbb{Z}/2\mathbb{Z})^3$. The only jumps allowed are those between x and $x \pm e_k$ where e_k is one of the three vectors of the canonical basis. The coefficients $Q(x, x')$ depend only on the classes modulo 2 of each coordinate of x and x' . A jump $\pm e_k$ changes the modulo class by 1 on the coordinate k . Once projected onto G_0 , the two jumps $x \pm e_k$ give the same transition on the cube. Q is then parametrized by 8×6 parameters (the cardinal of Q_0 times the number of directions).

In order to kill in a natural way the asymptotic drift v of the process, we assume a central symmetry such that $Q(x, x \pm e_k) = Q(x + (1, 1, 1), x + (1, 1, 1) \mp e_k)$. The model is thus parametrized by $24 = 8 \cdot 6/2$ parameters. In the generic case, the area anomaly Γ is non-zero.

Simulations are made for parameters chosen as in figure 5 with $u = 9/10$ and $v = 1/10$ run over 64.10^6 simulations and the process is observed at time $n = 40000$. We obtain the following values for both coordinates of $X_n^{(i)}/\sqrt{n}$:

- the empirical means are -0.0025 , -0.0020 and -0.0025 .
- the empirical covariance matrix has three coefficients 0.03001 on the diagonal and the other coefficients are all below 10^{-8} .
- the empirical third cumulants are all three below 10^{-6}
- the kurtosis are all three 3.0007 , 3.0006 and 3.0004 .

The empirical values for the area $A_n^{ij}/(\sigma_i \sigma_j n)$ normalized by the empirical standard deviations of the coordinates are:

- empirical mean $\Gamma^{12} = 1.500$, $\Gamma^{23} = 1.500$ and $\Gamma^{31} = -1.500$ (the area anomalies),

- empirical standard deviations are 0.5011, 0.5011 and 0.5011.
- empirical third cumulants $-1.46 \cdot 10^{-4}$, $-6.9 \cdot 10^{-5}$ and $1.1 \cdot 10^{-4}$
- empirical fourth cumulants 0.12533, 0.12535 and 0.12532.

All the cumulants correspond to a normal law for the coordinates and a Lévy drifted area for the area process.

5. APPENDIX

5.1. Rough paths. The theory of rough paths was introduced by Terry Lyons in the 90's. Over the years the theory has received important contributions from mathematicians like Peter Friz and Nicolas Victoir ([8]), Massimiliano Gubinelli ([9]), more recently Emmanuel Breuillard ([4] Antoine Lejay ([12]), Ismael Bailleul ([2]), Paul Gassiat etc. It has also inspired the more general theory of regularity structures developed by Martin Hairer in [16]. The quick overview of the theory we give here is based on the analytical approach developed by Terry Lyons and we leave out the algebraic point of view introduced by Massimiliano Gubinelli.

Suppose one has to give a meaning to the integral $\int_s^t g(X_r) dX_r$ or to the solution of a controlled differential equation of the type $dY_t = f(Y_t) dX_t$, $Y_0 = \xi(*)$, where X is some path in \mathbb{R}^d , called *driving signal*, and f, g some functions. The classical theory (for example Young's integral) fails whenever $(X_t)_{t \geq 0}$ is of low regularity (strictly smaller than $\frac{1}{2}$), even if the functions f and g are smooth. *The rough path theory* palliates this by lifting a partial differential equation to a space of paths that hold exhaustive information as driving signals, endowed with a topology that allows an important gain in regularity.

The idea of these lifted paths is inspired by the Taylor development for a function. For a path $x : [0, t] \rightarrow \mathbb{R}^d$ with bounded variations, we construct the *signature* of x as

$$S_{s,t}(x) = \left(1, \int_s^t dx_{s_1}, \dots, \iint_{s < s_1 < \dots < s_n < t} dx_{s_1} \otimes \dots \otimes dx_{s_n}, \dots \right)$$

where \otimes is the tensor product on \mathbb{R}^d . In the case of a path x whose regularity is too low to integrate against it, which is, after all, the one that matters most, we define each level of $S_{s,t}(x)$ as an abstract object with precise algebraic properties and satisfying some regularity conditions. The signature of a path is unique and encodes the path entirely.

Nonetheless, we do not need all the information contained in $S(x)$ to insure a good representation of solutions of differential equations: if $x : [0, t] \rightarrow \mathbb{R}^d$ is a path of regularity $\alpha \in (0, 1)$, we only need the informations contained in $S(x)$ up to level $\lfloor 1/\alpha \rfloor$. This new object obtained by truncating $S(x)$ is *the rough path* corresponding to x .

Example (Brownian motion rough path). Since the Brownian motion is a process of regularity $(1/2)^-$, it is sufficient to consider only the first three components of the signature. It can be shown that

$$\begin{aligned}\mathbb{B}_{s,t} &= \left(1, \int_s^t dB_u, \frac{1}{2}(B_t - B_s)^{\otimes 2} + A_{s,t}\right) \\ &= \left(1, \int_s^t dB_u, \iint_{s < u_1 < u_2 < t} \circ dB_{u_1} \circ dB_{u_2}\right)\end{aligned}$$

is the rough path for Brownian motion (see for example [15], chapter 3).

Let us now go back to the controlled differential equation (*). If we denote by \mathbb{X} the rough path corresponding to X , the solution of this new differential equation is also a rough path, which is denoted by \mathbb{Y} . In this way, if we consider (*) to be a differential equation in the Stratonovich sense, we get $d\mathbb{Y}_t = f(\mathbb{Y}_t)d\mathbb{X}_t(**)$. Since a rough path determines uniquely its corresponding path, the solution of (**) gives sense to and determines the solution of (*).

5.2. The group $G_2(E)$.

5.2.1. The general construction. In this section, we rewrite some results from the rough path theory from [8] (in particular from chapter 7) in order for them to correspond to the case of a finite-dimensional vector space E on \mathbb{R} . We concentrate on the case that is of interest to this article, namely $N = 2$. For more details and the general case $N \geq 2$ see [8] or [7].

We introduce the *tensorial truncated algebra* $T^{(2)}(E) = \bigoplus_{k=0}^2 E^{\otimes k}$, where \otimes is the tensorial product on E ($E^{\otimes 0} = \mathbb{R}$) and \bigoplus denotes a direct sum. Endowed with the operation

$$(13) \quad (a, b, c) \otimes_2 (x, y, z) = (ax, ay + bx, az + cx + b \otimes y)$$

it is a non-commutative algebra with unit element $(1, 0_E, 0_{E \otimes E})$.

For $x \in C^{1-var}([s, t], E)$ (the set of all continuous paths of finite 1-variation), the element that is given by

$$S_2(x)_{s,t} = \left(1, \int_{s < u < t} dx_u, \int_{s < u_1 < u_2 < t} dx_{u_1} \otimes dx_{u_2}\right) \in T^{(2)}(E)$$

is none other than the signature of x truncated at level 2. This object satisfies *Chen's relation*, i.e., for $0 \leq s < r < t \leq 1$

$$(14) \quad S_2(x)_{s,t} = S_2(x)_{s,r} \otimes_2 S_2(x)_{r,t}$$

and, in this particular case, \otimes_2 can be viewed as a path concatenation operator.

As in section 7.5.1 in [8], we define the set $G_2(E)$ by

$$G_2(E) = \{S_2(x)_{0,1} : x \in C^{1-var}([0, 1], E)\}$$

We now denote by $\mathbb{X} = (1, \mathbb{X}^{(1)}, \mathbb{X}^{(2)})$ an element of $G_2(E)$, where $\mathbb{X}^{(1)}$ stands for the first-order and $\mathbb{X}^{(2)}$ for the second-order increments. Implicitly, $\mathbb{X}_{s,t} = S_2(x)_{s,t}$ for some $x \in C^{1-var}([0, 1], E)$, and $\mathbb{X}_t = \mathbb{X}_{0,t}$. Since the symmetrical part of $\mathbb{X}^{(2)}$ depends on $\mathbb{X}^{(1)}$ (as $\int y dy = \frac{1}{2}y^2$), we can cut off redundant information by transforming $\mathbf{X}_t^{(2);i,j} = \int_0^t (X_s^i - X_0^i) dX_s^j - \int_0^t (X_s^j - X_0^j) dX_s^i$ for $1 \leq i, j \leq \dim(E)$. Under this new form, the element \mathbb{X} belongs to the space $\bigoplus_{k=0}^2 E^{\wedge k}$, where \wedge is

the antisymmetric tensor product on E : for $u, v \in E$, $u \wedge v = u \otimes v - v \otimes u$. For commodity reasons, we can use a more informal notation by neglecting the first component 1.

We finish this section with a particularly important example. Consider the space $H(\mathbb{R}^2) = \mathbb{R}^2 \oplus \mathbb{R}$, endowed with the following operation \star :

$$(x_1, x_2, a) \star (x'_1, x'_2, a') = \left(x_1 + x'_1, x_2 + x'_2, a + a' + \frac{1}{2}(x_1 x'_2 - x_2 x'_1) \right)$$

In this case, $H(\mathbb{R}^2)$ is a Lie group, called the *Heisenberg group* of dimension 3. Furthermore, $H(\mathbb{R}^2)$ can be identified with $G_2(\mathbb{R}^2)$.

5.2.2. The Carnot-Caratheodory distance on $G_2(E)$. It is natural to ask what is the shortest path in E for a given signature. The answer to this question allows to define the *Carnot-Caratheodory norm on $G_2(E)$* by

$$(15) \quad \|g\| := \inf \left\{ \int_0^1 |dx| : x \in C^{1-var}([0, 1], E) \text{ and } S_2(x)_{0,1} = g \right\}$$

where $|\cdot|_E$ is a restriction to E of the Euclidean norm.

Since the norm thus defined is homogeneous ($\|\delta_\lambda g\| = |\lambda| \|g\|$ for $\lambda \in \mathbb{R}$), symmetric ($\|g\| = \|g^{-1}\|$) and sub-additive ($\|g \otimes h\| \leq \|g\| + \|h\|$), it induces a left-invariant, continuous metric \mathbf{d} on $G_2(E)$ through the application

$$(16) \quad \begin{array}{ccc} \mathbf{d} : G_2(E) \times G_2(E) & \rightarrow & \mathbb{R}_+ \\ (g, h) & \mapsto & \|g^{-1} \otimes_2 h\| \end{array}$$

In this case, $(G_2(E), \mathbf{d})$ is a geodesic space (in the sense of definition 5.19 from [8]). It is also a Polish space (corollary 7.50 from [8]).

The Carnot-Caratheodory norm is difficult to use for practical estimations but we can give it a good upper bound:

Proposition 5.1. *There exists $\nu > 0$ such that, for \mathbf{d} defined as above, for any $\mathbb{X} \in G_2(E)$ and $0 \leq s < t \leq 1$, we have:*

$$(17) \quad \mathbf{d}(\mathbb{X}_s, \mathbb{X}_t) = \|\mathbb{X}_{s,t}\| \leq \nu \left(\left| \mathbf{X}_{s,t}^{(1)} \right|_E + \left| \mathbf{X}_{s,t}^{(2)} \right|_{E \otimes E}^{\frac{1}{2}} \right)$$

5.3. The signed area.

5.3.1. Definition and properties. If $(x_t, y_t)_{t \geq 0}$ is a smooth curve in \mathbb{R}^2 , the signed area associated to it is the real-valued function given by $(s, t) \mapsto a(x, y)_{s,t} = \frac{1}{2} \int_s^t x_u dy_u - \int_s^t y_u dx_u$. The following lemma gives a first description of a through its properties:

Lemma 5.1. *The signed area has the following important properties:*

- 1 *Invariance by translation of (x, y) by a constant vector: $a((x, y) + (v_1, v_2)) = a(x, y)$*
- 2 *Invariance by rotation of (x, y) by an angle $\theta \in]0, 2\pi[$: $a(\theta(x, y)) = a(x, y)$*
- 3 *Scaling by dilatation of (x, y) by $c > 0$: $a(c(x, y)) = a(x, y)$*
- 4 *Symmetry property: given (x, y) and (x', y') two symmetrical paths with respect to an axis, $a(x, y) = -a(x', y')$*

$$5 \text{ Law of concatenation: } \forall s \leq v \leq t, \ a_{s,t} = a_{s,v} + a_{v,t} + \frac{1}{2}(x_v - x_s)(y_t - y_v) - (y_v - y_s)(x_t - x_v)$$

This last very important property leads us to a geometric interpretation of the group $G_2(\mathbb{R}^2)$: the operation \star we have endowed it with can now be seen as the concatenation of two bidimensional curves and their signed areas.

5.3.2. *Lévy area and Brownian motion on Heisenberg group.* We now come to an object that plays a central part in this article,

Definition 5.1. *Let $B = (B^1, B^2)$ be standard two-dimensional Brownian motion. The signed area between the curve of B on the interval $[s, t]$ and the segment $[B_s, B_t]$ is called Lévy area and is given by*

$$\mathcal{A}_{s,t} = \frac{1}{2} \iint_{s < u_1 < u_2 < t} dB_{u_1}^1 dB_{u_2}^2 - dB_{u_1}^2 dB_{u_2}^1$$

The law of the Lévy area is given by

$$\mathbb{E} [e^{\theta \mathcal{A}_1} | B_1 = z] = \frac{\theta}{\sinh \theta} e^{-\frac{|z|^2}{2}} (\theta \coth \theta - 1)$$

as computed by Paul Lévy or, more recently, by Daniel Levin and Mark Wildon in [14] using shuffle products. When the signed area presents an additional drift γ , the above formula becomes:

$$\mathbb{E} [e^{\theta(\mathcal{A}_1 + \gamma)} | B_1 = z] = e^{\theta \gamma} \frac{\theta}{\sinh \theta} e^{-\frac{|z|^2}{2}} (\theta \coth \theta - 1)$$

The Lévy area plays a crucial role in the construction of the Brownian motion on the Heisenberg group: it is then the process $(B_t, \mathcal{A}_t)_{t \geq 0}$. From the example 5.1, we can also see that the second-level component of the Brownian motion rough path is determined by its Lévy area, and, moreover, there is an obvious correspondence between the rough path representation of a standard Brownian motion and the Brownian motion on Heisenberg group.

6. OPEN QUESTIONS

The present result leads to some open questions both about the limit process with the area anomaly and about the discrete models which may converge to such a limit.

It would be interesting to understand how the area anomaly fits in the Fock space description of Brownian Motion: the question is non-trivial because the Lévy area leaves in the second chaos and the presence of an area anomaly adds a zero-chaos component to the Lévy area.

Next, two-dimensional Brownian motion is known to exhibit conformal symmetry and it is natural to ask how our limit process behaves under conformal transformations.

Since we focus on the area drift, an important question that arises is: can Girsanov's theorem be extended to cancel the area anomaly by a change of measure on the rough path space?

Our proof exhibits striking similarities with [1]. Our "internal" G_0 space seems to play the same role as the compact sphere in their paper and their proof also uses

theory of rough paths to control convergences. We would like to know if one can build models on Riemannian manifolds which may exhibit area anomalies.

Finally, Brownian motion belongs to the larger family of Lévy processes and, consequently, we may wonder whether Lévy processes may be also enhanced with area components and approximated by suitable discrete processes.

REFERENCES

- [1] J. Angst, I. Bailleul, and C. Tardif. Kinetic brownian motion on riemannian manifolds. *Electronic Journal of Probability*, 20:1–40, 2015.
- [2] I. Bailleul. Flows driven by rough paths. *Revista Iberoam. Math.*, 31(3):29, 2015.
- [3] Erich Baur. An invariance principle for a class of non-ballistic random walks in random environment. *Preprint arxiv:math/1408.1690*, 2014.
- [4] E. Breuillard, P. Friz, and M. Huesmann. From random walks to rough paths. *Proc. Amer. Math. Soc.*, 137(10):3487–3496, 2009.
- [5] N. Enriquez and Y. Kifer. Markov chains on graphs and brownian motion. *Journal of Theoretical Probability*, 14(2), 2001.
- [6] P. Friz, P. Gassiat, and T. Lyons. Physical brownian motion in a magnetic field as a rough path. *Trans. Amer. Math. Soc.*, 367:7939–7955, 2015.
- [7] P. K. Friz and M. Hairer. *A Course on Rough Paths - With an introduction to regularity structures*. Springer, 2014.
- [8] P. K. Friz and N. B. Victoir. *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge University Press, 2010.
- [9] M. Gubinelli. Controlling rough paths. *J. Func. Anal.*, 216:86–140, 2004.
- [10] T. Kazami and K. Uchiyama. Random walks on periodic graphs. *Transactions of the American Mathematical Society*, 360(11):6065–6087, 2008.
- [11] W.S. Kendall. Coupling all the lévy stochastic areas of multidimensional brownian motion. *Ann. Probab.*, 35(3):935–953, 2007.
- [12] A. Lejay. Introduction to rough paths. *Séminaire de Probabilités XXXVII. Lecture Notes in Math*, 1832:1–59, 2003.
- [13] A. Lejay and T. J. Lyons. On the importance of the levy area for studying the limits of functions of converging stochastic processes. application to homogenization. *Current Trends in Potential Theory*, page 63–84, 2003.
- [14] D. Levin and M. Wildon. A combinatorial method for calculating the moments of levy area. *Trans. Amer. Math. Soc.*, 360(12), 2007.
- [15] T. J. Lyons, M. Caruana, and Th. Lévy. *Differential Equations Driven by Rough Paths (Ecole d’Eté de Probabilités de Saint-Flour XXXIV)*. Springer, 2004.
- [16] M. Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2):269–504, 2014.
- [17] D. W. Stroock and S. R. S. Varadhan. Limit theorems for random walks on Lie groups. *Sankhyā Ser. A*, 35(3):277–294, 1973.

E-mail address: olga.lopusanschi@upmc.fr

E-mail address: damien.simon@upmc.fr

LABORATOIRE DE PROBABILITÉS ET MODÈLES ALÉATOIRES (LPMA), UPMC-SORBONNE UNIVERSITÉ, UMR CNRS 7599, CASE 188, 4, PL. JUSSIEU, F-75252 PARIS CEDEX 5, FRANCE.